VIROLOGY 224, 498–509 (1996)
ARTICLE NO. 0556

Reciprocal Effects of Splicing and Polyadenylation on Human Immunodeficiency

View metadata, citation and similar papers at core.ac.uk
brought
provided by Elsevier

JEANNINE M. SCOTT* and MICHAEL J. IMPERIALE*,†,[1]

*Department of Microbiology and Immunology and †Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, Michigan 48109-0620

Received June 21, 1996; accepted August 14, 1996

Insertion of a functional splicing cassette into a construct containing the HIV-1 poly(A) site followed by the adenovirus L3 poly(A) site results in both specific stimulation of 3′ end processing at the HIV-1 site and an increase in the steady-state levels of RNA processed at both sites. To further evaluate this influence of splicing on processing of the HIV-1 poly(A) site, defined mutations which alter or abolish splicing of the intron were made and analyzed for their effects on polyadenylation and steady-state levels of RNA. The data show that a point mutation at the 3′ splice site caused activation of a cryptic splice acceptor that is as efficient as the wild-type acceptor. Substitution of this mutant intron for the wild-type intron resulted in stimulation of the HIV-1 poly(A) site to levels equivalent to those caused by the wild-type intron. This mutant did not, however, have as great an effect on steady-state RNA levels as the wild-type intron. A second construct containing a mutated branch point and polypyrimidine tract resulted in abolishment of splicing and a decrease in both poly(A) site use and steady-state levels of RNA. These data demonstrate that the enhanced use of the HIV-1 poly(A) site is a direct result of the splicing reaction, and not merely due to the sequences that were inserted. The effect that poly(A) site strength has on splicing was also addressed. Using activation of the cryptic splice acceptor to indicate changes in splicing efficiency resulting from alterations in poly(A) site strength, it was determined that poly(A) site strength does have an effect on the efficiency of the splicing reaction. © 1996 Academic Press, Inc.

INTRODUCTION

The biogenesis of mRNA in higher eukaryotes involves substantial processing of primary transcripts to produce mature mRNA molecules. Two of these processing steps are polyadenylation, which forms the 3′ end, and splicing, which removes intervening sequences. Much is known about the *cis*-acting sequences and *trans*-acting factors that are essential for these reactions. Polyadenylation is the site-specific endonucleolytic cleavage of the pre-mRNA followed by the template-independent addition of approximately 200 adenosine residues which make up the poly(A) tail (for reviews, see Manley, 1988; Wahle and Keller, 1992; Wahle, 1995). The minimal sequences required for polyadenylation are the highly conserved hexanucleotide sequence, AAUAAA, that lies 10–30 nt upstream of the poly(A) addition site and a G/U- or U-rich sequence found within 50 nt downstream of the cleavage site. Additional accessory sequences that can regulate the efficiency of poly(A) site use have also been identified both upstream and downstream of the core signals in several viral systems (Dougherty and Temin, 1987; Carswell and Alwine, 1989; DeZazzo and Imperiale, 1989; Russnak and Ganem, 1990; Sanfacon *et al.,* 1991; DeZazzo *et al.,* 1991; Valsamakis *et al.,* 1991; Brown *et al.,* 1991). Several multicomponent protein factors are essential for the cleavage and polyadenylation reaction and include cleavage/polyadenylation specificity factor (CPSF), cleavage stimulatory factor (CstF), poly(A) polymerase (PAP), and cleavage factors I and II (CFI and CFII).

Pre-mRNA splicing is the accurate removal of introns and joining of exons. Similar to polyadenylation, splicing also involves specific sequences and protein factors (for reviews, see Green, 1991; Moore *et al.,* 1993). The consensus sequences necessary for splicing are a GU dinucleotide at the 5′ splice site and an AG dinucleotide at the 3′ splice site, along with the branch point residue and polypyrimidine tract located just upstream of the 3′ splice site. The length and pyrimidine content of the polypyrimidine tract have been shown to be key determinants in the splicing strength of an intron (Reed, 1989; Roscigno *et al.,* 1993). Careful analysis of the sequences surrounding splice sites has revealed that other sequences located in both introns and exons can also influence the efficiency of splicing (Nesic and Maquat, 1994; Lavigueur *et al.,* 1993; Humphrey *et al.,* 1995; Tian and Maniatis, 1993, 1994; Watakabe *et al.,* 1993). The splicing reaction is catalyzed by a large multicomponent complex

[1] To whom correspondence and reprint requests should be addressed at University of Michigan Medical School, 6714 Medical Science II, Ann Arbor, MI 48109-0620. Fax: (313) 764-3562. E-mail: imperial@umich.edu.

0042-6822/96 $18.00
Copyright © 1996 by Academic Press, Inc.
All rights of reproduction in any form reserved.

called the spliceosome, which includes small nuclear ribonucleoprotein particles (snRNPs) and additional essential proteins. Much of our understanding of the mechanisms of polyadenylation and splicing has come from cell-free systems in which the two events are studied separately; a poly(A) site can be cleaved in the absence of splicing and a pre-mRNA which contains an intron but lacks a poly(A) site can still be spliced. Additionally, the factors which have been shown to be essential for polyadenylation and splicing have not shown any overlap. In spite of this, the observation that there are no examples of mRNAs that are spliced without also being polyadenylated suggests that the two processing steps may be linked in the cell.

The idea that splicing and polyadenylation may be coupled reactions is consistent with the exon definition theory of splice site recognition (Robberson *et al.,* 1990), which states that a 3′ terminal exon is identified by splicing factors binding at its 5′ end and by polyadenylation factors binding at its 3′ end. Indeed, there have been several reports to date that demonstrate coupling of the splicing and polyadenylation reactions. Berget and colleagues have shown *in vitro* that polyadenylation can be stimulated by an intron located upstream in the precursor RNA (Niwa *et al.,* 1990). More recently, they have shown that polyadenylation can be activated *in vitro* by a downstream intron enhancer containing a polypyrimidine tract and a pseudo-5′ splice site sequence (Lou *et al.,* 1996). This enhancer was shown to bind to several splicing factors and the binding of these factors was required for activation of 3′ end cleavage, providing further evidence that splicing factors can influence polyadenylation *in vitro.*

Several transfection systems have also demonstrated that functional introns can have a profound effect on polyadenylation and subsequent gene expression (Chiou *et al.,* 1991; Miller and Stoltzfus, 1992; Huang and Gorman, 1990; Villarreal and White, 1983; Nesic and Maquat, 1994; Nesic *et al.,* 1993; DeZazzo *et al.,* 1992; Liu and Mertz, 1993). This positive influence of splicing on polyadenylation may be due, at least in part, to the involvement of the U1 snRNP-A protein (U1A), which has been shown to bind to the upstream efficiency element in the simian virus 40 late poly(A) site (Lutz and Alwine, 1994; Wassarman and Steitz, 1993). The U1A protein can interact with both the U1 snRNA and the pre-mRNA simultaneously via its two RNA recognition motifs (Lutz and Alwine, 1994), thereby potentially serving as a bridging molecule for communication between the splicing and polyadenylation machineries. Indeed, U1A has recently been shown to contact the 160-kDa subunit of CPSF directly (Lutz *et al.,* 1996).

With respect to the reciprocal influence polyadenylation may have on splicing, we are aware of only two reports, the data of which point to conflicting conclusions. *In vitro* studies suggested that polyadenylation can positively influence splicing in that, when 3′ end processing signals were mutated, splicing of the proximal upstream intron was depressed without affecting splicing of more distant upstream introns (Niwa and Berget, 1991). The conflicting report looked at the effect of polyadenylation on splicing in the human triosephosphate isomerase gene in intact cells and indicated that the efficiency of the poly(A) site had no effect on splicing of the two nearest introns (Nesic *et al.,* 1995). Therefore, this debate regarding the reciprocal effects of the processing events is not closed.

Viruses, because of their ability to usurp the cell's machinery for their own gene expression and regulation, can provide powerful systems for exploring the various echelons of control that the cell has developed. Previous studies from our laboratory which addressed the relative roles of promoter proximity and upstream sequences in HIV-1 mRNA 3′ end formation demonstrated that a functional splicing cassette inserted between the promoter and the poly(A) site could specifically increase HIV-1 poly(A) site use as well as the total levels of processed transcripts in the steady-state pool (DeZazzo *et al.,* 1992). In these studies, we deleted the 5′ or the 3′ half of the intron to determine if these effects were due to the sequences surrounding either of the splice sites individually. Analysis of the processing of transcripts produced by these constructs showed that the 5′ splice site alone had no effect on steady-state RNA levels. The 3′ splice site alone, however, caused an increase in the steady-state levels of RNA and specifically enhanced the use of the HIV-1 poly(A) site. Further analysis showed that, even though the 5′ splice donor sequences had been deleted, a significant number of these transcripts were being spliced to a cryptic 5′ splice site and the increase in the steady-state level of polyadenylated RNA correlated with the observed level of cryptic splicing. The experiments presented in this report were designed to investigate further the apparent coupling of splicing and polyadenylation with respect to the strength of the HIV-1 poly(A) site. We have made defined mutations in the intron that impair or abolish splicing and evaluated the effect that these mutations have on both processing of the HIV-1 poly(A) site and steady-state RNA levels. Our results show that when mutations are made in the intron that render it nonfunctional, there is no enhancement of 3′ end processing nor any stimulation of steady-state RNA levels, thus providing direct evidence of a positive role for splicing on polyadenylation. We have also examined the effects polyadenylation has on the splicing efficiency of our intron-containing constructs. Using two versions of the HIV-1 poly(A) site, either containing long terminal repeat (LTR) U3 sequences, which make the site very efficient, or lacking these sequences, we find that poly(A) site strength does influence the splicing reaction. These

results are consistent with a model in which the polyadenylation and splicing machineries positively influence each other in order to define and process the terminal exon of a pre-mRNA molecule. Within the compact genome of HIV-1, there are several alternative splice sites which are used to produce at least seven different transcripts, and this level of complexity suggests a sophisticated method of regulating gene expression. The experiments in this report explore the possible interactions between polyadenylation and splicing that accomplish this intricate regulation of gene expression.

EXPERIMENTAL PROCEDURES

Plasmid constructions

The plasmids used in this study are derivatives of pGEM 3Zf- (Promega) containing a 435-bp adenovirus major late promoter (MLP) fragment followed by tandem (Fig. 1A) or single (Fig. 3A) poly(A) sites. Tandem poly(A) site constructs contain a 400-bp fragment of the HIV-1 LTR followed directly by the adenovirus L3 poly(A) site and have been described previously (DeZazzo *et al.,* 1992). Single poly(A) site constructs were made which contain the HIV-1 poly(A) site fragment with or without LTR U3 sequences. The construct lacking U3 (ΔU3) has an equivalent size fragment from the HIV-1 envelope gene which has no effect on polyadenylation (DeZazzo *et al.,* 1991). The splicing cassette used for these constructs is a 120-nucleotide intron derived from adenovirus 2 sequences (Niwa *et al.,* 1990) and in this paper is referred to as "wild type." To make the construct containing a 3′ splice site mutation, the intron from pMXSVL 3′ (kindly provided by Sue Berget (Niwa *et al.,* 1990)) was amplified by PCR using exon-specific primers. This intron has a single base change (G → A) at the consensus splice acceptor dinucleotide (Fig. 1A). A second splicing mutant was made by PCR site-directed mutagenesis to inactivate the branch point and polypyrimidine tract. Exact base changes in the mutant are indicated in Fig. 1A. The mutant introns were then inserted between the MLP sequences and the poly(A) sites via *Bam*HI linkages. Mutations were confirmed by dideoxynucleotide sequencing. Other plasmids used in this study include pU3F-L3, pU3C-L3, p3′U3-L3, and pΔBSVF and have been described previously (DeZazzo *et al.,* 1992). pΔBSVF contains the MLP followed by 160 nucleotides of U3-R sequences from HIV-1 and the SV40 early poly(A) site and is used to control for differences in transfection efficiency. Plasmids were propagated in *Escherichia coli* DH5α and purified using the Qiagen plasmid purification protocol.

Transient transfections and RNA preparation

The human 293 cell line (Graham *et al.,* 1977) was maintained in Dulbecco's modified Eagle's medium (Gibco-BRL) supplemented with 10% fetal bovine serum (Hyclone Laboratories, Inc.) and 0.2 m*M* L-glutamine (Gibco-BRL). Monolayer cultures at approximately 40% confluence were transfected with 10 μg assay plasmid and 5 μg control plasmid by the $CaPO_4$ precipitation method. The DNA was removed 24 hr after transfection, and RNA was purified 24 hr later by acid guanidinium isothiocyanate–phenol–chloroform extraction (Chomczynski and Sacchi, 1987). The RNA was quantitated by A_{260} spectroscopy and agarose gel electrophoresis. The poly(A)$^+$ fraction was isolated from 50 μg total RNA using oligo(dT)-cellulose chromatography (Boehringer Mannheim Biochemicals). The poly (A)$^-$ RNA fraction was obtained by depleting total RNA of poly(A)$^+$ RNA via sequential oligo(dT) columns until we could no longer detect the poly(A)$^+$ transfection control RNA by S1 nuclease protection.

S1 nuclease mapping and quantitation

To determine the steady-state levels of transcripts polyadenylated at the HIV-1 or L3 poly(A) sites, poly(A)$^+$ RNA fractions were coprecipitated with 30 ng 3′ end-labeled probe (5×10^5 to 1×10^6 cpm) and 25 μg yeast tRNA. The probe was made by filling in the ends of an 816-bp *Xba*I fragment that spans both poly(A) sites with [α-^{32}P]dCTP (Amersham) using Klenow enzyme (BRL). The levels of spliced product from constructs containing splicing signals were determined using 5′ end-labeled probes. These probes were made by linearizing plasmid DNA (pIVSU3-L3, p3′MU3-L3, or pPTMU3-L3) with *Bgl*II, which cuts in the HIV-1 R sequence, labeling the 5′ end with [γ-^{32}P]dATP using T4 polynucleotide kinase, and finally digesting upstream of the transcription start site with *Xho*I for constructs containing U3 sequences, or with *Stu*I for ΔU3 constructs. In each case, a labeled fragment of 680 nucleotides is produced which spans the promoter and splice signals. These probes detect both spliced and full-length assay transcripts as well as the control RNA. The poly(A)$^+$ RNA fractions were hybridized overnight at 52° for the 3′ end-labeled probe, or 57° for the 5′ end-labeled probe. Hybridized samples were then treated with S1 nuclease (BRL) at 25° for 2 hr under conditions that have been previously described (DeZazzo *et al.,* 1992). The protected fragments were resolved on a 6% polyacrylamide–8 *M* urea gel and visualized by autoradiography. The gels were also exposed to a PhosphorImager screen for quantitative analysis using ImageQuant software (Molecular Dynamics). These values were used to determine the total levels of polyadenylated RNA or the percentage of spliced product for each construct.

RESULTS

Our previous work which addressed the relative roles of promoter proximity and upstream sequences on pro-

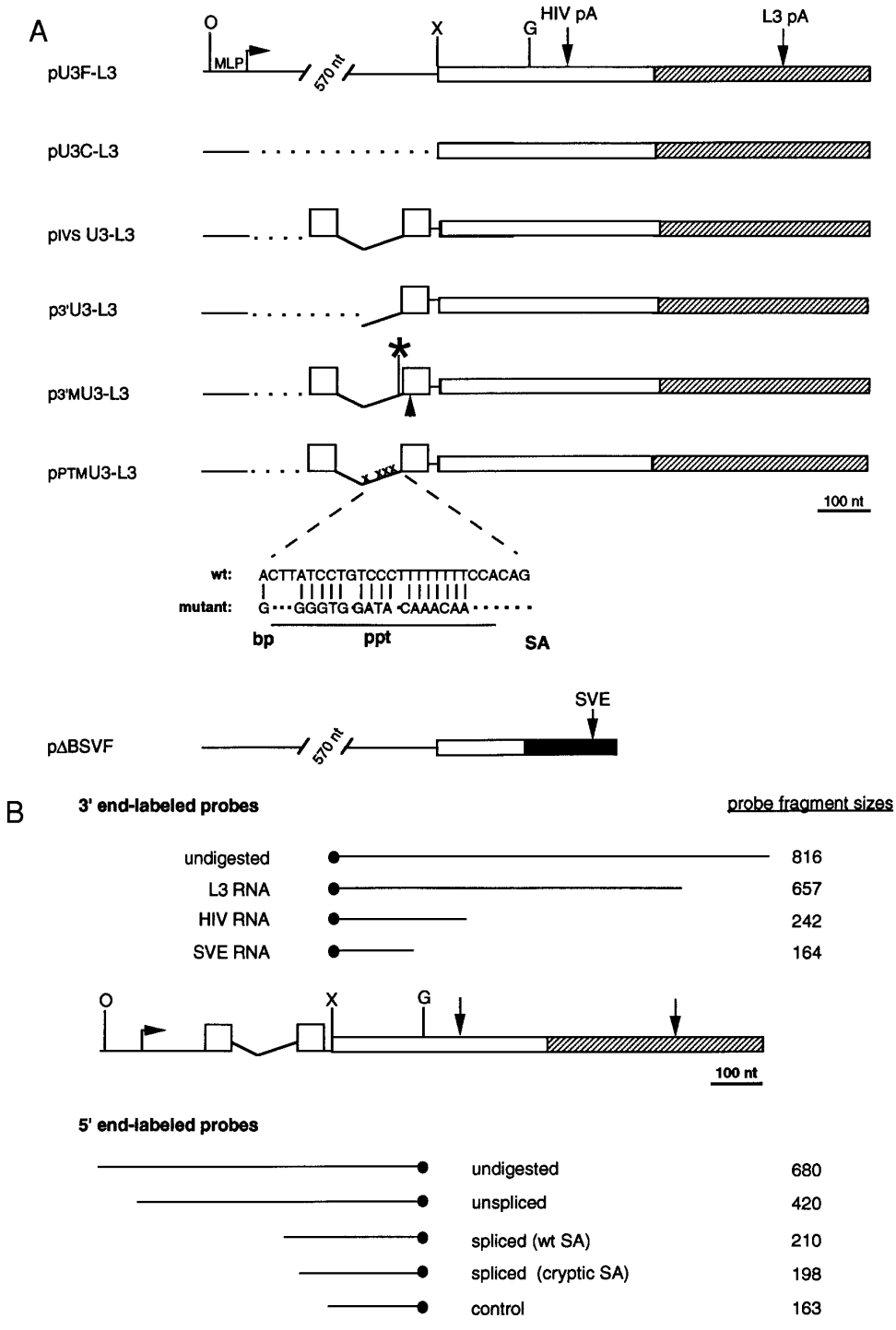

FIG. 1. Diagram of transfection constructs and predicted S1 analysis products. (A) Diagram of constructs. Each construct contains two poly(A) sites: the HIV-1 poly(A) site including stimulatory U3 sequences (open boxes) and the adenovirus L3 poly(A) site (hatched boxes). Cleavage sites are indicated by the downward arrows. Rightward arrow denotes the transcription start site of the adenovirus major late promoter (MLP). Constructs containing wild-type and mutant splicing cassettes inserted into pU3C-L3 are also diagrammed. The splicing cassette is designated by the large open boxes connected by the V-shaped line. In p3′U3-L3, the splice donor and 5′ half of the intron have been deleted. The large asterisk in p3′MU3-L3 represents the splice acceptor (SA) point mutation AG → AA and the upward arrowhead marks the cryptic SA that is activated (see text). The specific mutations made at the branch point (bp) and within the polypyrimidine tract (ppt) to create pPTMU3-L3 are indicated. Transfection control plasmid pΔBSVF contains HIV-1 U3 sequences followed by the SV40 early poly(A) site (SVE). Restriction sites: O, *Xho*I; X, *Xba*I; G, *Bgl*II. (B) S1 probes and predicted fragments to detect polyadenylation (3′ end-labeled at the *Xba*I site) or splicing (5′ end-labeled at the *Bgl*II site). Probe and protected fragment sizes from the various RNA species are indicated. A simplified diagram of the constructs is provided for reference.

cessing of the HIV-1 poly(A) site demonstrated that when a functional splicing cassette was inserted into constructs which contained the complete HIV-1 poly(A) site (i.e., with LTR U3 sequences), both the strength of this site relative to a downstream site and the steady-state level of processed RNA were enhanced (DeZazzo *et al.,* 1992). The positive influence of splicing on polyadenylation was supported by experiments in which a splicing cassette containing only the 3′ half of the intron also stimulated polyadenylation, albeit at an intermediate level. This degree of stimulation correlated with an intermediate level of splicing compared to that of the wild-type intron and was most likely due to the lack of an authentic splice donor and consequent use of a cryptic donor site(s). Based on these data, we hypothesized that the increases in poly(A) site use and in steady-state levels of RNA were a direct result of the splicing reaction and not simply due to sequences that had been inserted.

Direct effects of splicing on polyadenylation

Others have shown that gene expression can be positively affected by the insertion of a functional intron into the expression construct (Buchman and Berg, 1988; Ryu and Mertz, 1989) and in some cases this enhancement is due to increased polyadenylation (Chiou *et al.,* 1991; Huang and Gorman, 1990). We wished to determine if the stimulation seen in our system was a direct result of splicing, or was instead simply due to some other form of enhancement by the inserted sequences. To this end, defined mutations were made in the splicing cassette that either impaired or abolished splicing. The transcripts from these constructs were evaluated with respect to the relative use of the HIV-1 poly(A) site, steady-state levels of RNA, and splicing efficiency. If the splicing reaction can indeed stimulate polyadenylation, alteration of the authentic splicing signals to suboptimal or nonfunctional sequences should cause a decrease in polyadenylated transcripts in the steady-state pool. The constructs used for this study are diagrammed in Fig. 1. Each construct contains the adenovirus MLP followed by tandem poly(A) sites, HIV-1 and adenovirus L3. The L3 poly(A) site serves as a default site for 3′ end processing of transcripts that are not processed at the HIV-1 site. The constructs used in our original study are pU3F-L3, pU3C-L3, and pIVSU3-L3. pU3F-L3 has the HIV-1 poly(A) site 1207 nt from the cap site. In pU3C-L3 the HIV-1 poly(A) site is located closer to the cap site (270 nt), and pIVSU3-L3 is pU3C-L3 with a 120-nt intron inserted between the MLP and the HIV-1 sequences. To make p3′MU3-L3, we amplified the intron from pMXSVL 3′ by PCR. This intron is identical to that found in pIVSU3-L3 except it contains a point mutation at the splice acceptor site to change the consensus AG to AA. *In vitro,* this mutation has been shown to inhibit splicing completely (Niwa *et al.,* 1990). An additional construct, pPTMU3-L3, was made by PCR site-directed mutagenesis. In this construct the branch point residue is changed from A to G and extensive purine substitutions are present in the polypyrimidine tract.

TABLE 1

Level of 3′ End Processing

Construct	HIV/L3[a]	HIV + L3/control[b]
pU3F-L3	34 ± 4	1.5 ± 0.1
pIVSU3-L3	39 ± 7	1.3 ± 0.3
p3′MU3-L3	30 ± 3	0.3 ± 0.1
p3′U3-L3	15 ± 5	0.20 ± 0.03
pU3C-L3	5 ± 1	0.05 ± 0
pPTMU3-L3	1.8 ± 0.2	0.03 ± 0.01

Note. Results are averages of 2–4 experiments ± standard deviations.

[a] The ratio of RNA processed at the HIV poly(A) site to that processed at the L3 poly(A) site.

[b] The total amount of RNA processed at both sites, normalized to control.

The plasmids were transfected side by side into the human embryonic kidney cell line 293. 293 cells constitutively express adenovirus E1A proteins, which transactivate the adenovirus MLP contained in the plasmids. Each plasmid was cotransfected with pΔBSVF as a control for transfection efficiency. The RNA generated from this control plasmid is used to calculate the relative levels of spliced or polyadenylated RNA, enabling us to compare the processing events on various transcripts that are assayed with different probes. We first examined 3′ end formation by S1 nuclease analysis of poly(A)$^+$ RNA using a 3′ end-labeled *Xba*I fragment that spanned both poly(A) sites (Fig. 2A and Table 1). Compared to pU3C-L3, the relative use of the HIV-1 poly(A) site is increased in pre-mRNAs containing the wild-type or the 3′ mutant intron (Fig. 2A, compare lanes 3 and 4 to lane 5). Additionally, steady-state levels of the processed transcripts are increased, although to a lesser degree for p3′MU3-L3. In contrast, however, there is no stimulation of processing and actually a reduction in steady-state RNA levels when the branch point and polypyrimidine tract are mutated (lane 6). Thus, it appears that specific mutations in critical splicing sequences adversely affect 3′ end processing.

To evaluate splicing in these transcripts, poly(A)$^+$ RNA was hybridized to 5′ end-labeled probes and treated with S1 nuclease (Fig. 2B). These probes include all intron sequences and therefore can detect spliced and unspliced products regardless of which poly(A) site is used. A probe made from pIVSU3-L3 was used for analysis of the RNA from the transfections with pIVSU3-L3 and p3′U3-L3. p3′U3-L3 contains only the 3′ half of the intron, resulting in a moderate level of stimulation (Table 1) that correlates with the intermediate level of splicing shown

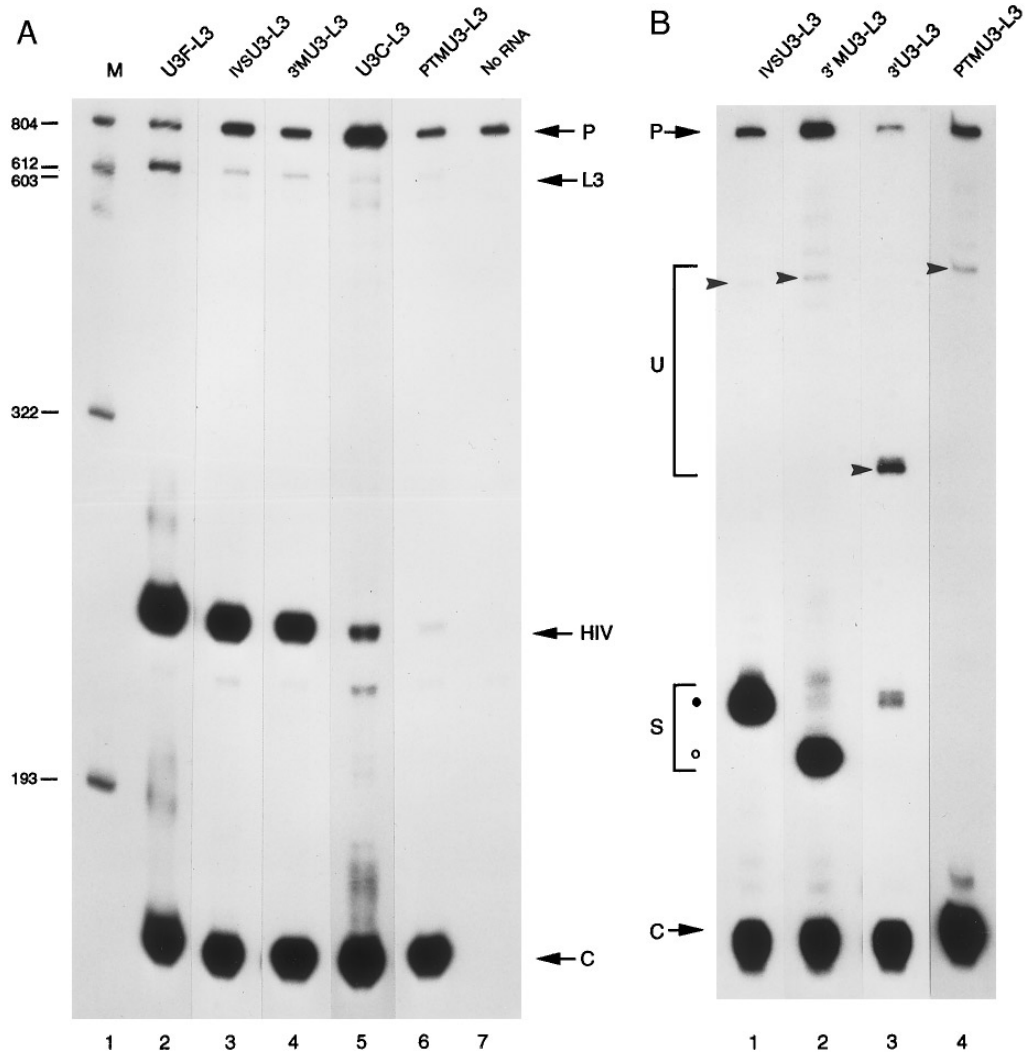

FIG. 2. S1 analysis of RNA processing events in the presence of wild-type and mutant splicing cassettes. (A) S1 analysis of 3′ end formation. S1 products from analysis of the poly(A)$^+$ fraction of 50 μg RNA are shown. Transcripts polyadenylated at the HIV-1 site protect a 242-nt probe fragment while transcripts polyadenylated at the L3 site result in a 657-nt fragment. The full-length probe (P) and the fragment protected by the transfection control RNA (C) are indicated. Lane 1, molecular weight markers, lanes 2–6, S1 analysis of the indicated transfection RNAs, lane 7, S1 sample containing only probe. (B) S1 analysis of splicing levels in poly(A)$^+$ fraction. Unspliced transcripts protect a fragment of 420-nt (arrowhead) while spliced transcripts protect 210 nt when the wild-type SA is used (filled circle), or 198 nt when the cryptic SA is used (open circle). The unspliced fragment size for p3′U3-L3 is shorter, 330 nt, because of deletion of the 5′ half of the intron.

in Fig. 2B, lane 3. To assay the splicing levels from p3′MU3-L3 and pPTMU3-L3, probes were made from the corresponding plasmids. Note that the spliced transcript from p3′MU3-L3 protects a shorter probe fragment than the wild-type spliced transcript, indicating that a cryptic splice acceptor has been activated. Probe fragments that migrate slower than the unspliced bands are background bands as judged by their presence in samples lacking RNA (data not shown). The radioactivity in each band was quantitated and used to determine the relative splicing efficiencies. In p3′MU3-L3, splicing to the cryptic acceptor was as efficient as that to the wild-type acceptor in the wild-type construct (98.4 vs 99.6%). In contrast, splicing was completely inhibited by the branch point–polypyrimidine tract mutations. Again, the level of poly(A) site stimulation shown in Fig. 2A correlates with these splicing activities, and we therefore conclude that the enhancement of the HIV-1 poly(A) site and of steady-state levels of RNA occurs as a result of splicing and is not due to nonspecific effects of sequences within the intron cassette.

Influences of poly(A) site strength on splicing

Having shown that splicing stimulates both 3′ end processing efficiency and steady-state RNA levels, we were next interested in determining if the opposite was true, i.e., whether poly(A) site strength has any influence on

splicing in our system. To measure effects on splicing efficiency due to changes in the strength of the poly(A) site, we used activation of the cryptic splice acceptor because we felt it would be very sensitive to perturbations of any potential interactions between the processing machineries. For these studies, we constructed plasmids that contain either the wild-type or the 3′ mutant intron followed by the HIV-1 poly(A) site (Fig. 3A). The strength of the poly(A) sites in these constructs was categorized as either "strong" or "weak" based on the presence or absence of the stimulatory U3 sequences, respectively (DeZazzo *et al.,* 1991, 1992). These constructs do not contain the very efficient L3 poly(A) site in order to avoid dominant effects that it may have on splicing when the weak HIV-1 poly(A) site is present. Within each pair of plasmids, the degree of splicing that occurs at the cryptic splice acceptor of the 3′ mutant intron indicates the effect that poly(A) site strength has on splicing: if poly(A) site strength can influence splicing, the strong poly(A) site should drive the splicing reaction to take place at the cryptic splice acceptor in the mutant intron, whereas the weak poly(A) site should result in reduced splicing at the cryptic site.

S1 analysis of the RNA with a 3′ end-labeled probe to detect polyadenylation demonstrates that, although both the strong and weak HIV-1 poly(A) sites are functional, the ΔU3 constructs produce much less steady-state RNA (Fig. 3B). A 5′ end-labeled probe used to analyze splicing shows that transcripts containing the strong poly(A) site with either intron were spliced very efficiently (Fig. 3C, lanes 1 and 2, Table 2). Additionally, the cryptic splice acceptor in the 3′ mutant intron is activated in the presence of either the strong or the weak HIV-1 poly(A) site (lanes 2 and 4). Quantitation of the bands revealed, however, that in the weak poly(A) site constructs, 95% of the poly(A)$^+$ transcripts containing the wild-type intron were spliced while only ~70% of the transcripts containing the 3′ mutant intron were spliced at the cryptic splice acceptor (Table 2). This shows that the strength of the poly(A) site does influence splicing in these transcripts. It was possible, however, that by examining only poly(A)$^+$ RNA, we might not have detected splicing events that occurred on poly(A)$^-$ transcripts. To examine whether a significant number of transcripts from p3′MΔU3 were being spliced but not polyadenylated, we analyzed the poly(A)$^-$ fraction (Fig. 4). Total RNA was fractionated using oligo(dT)-cellulose until the poly(A)$^+$ RNA signal, as measured by the 5′ end-labeled probe, was no longer detectable in the eluates by S1 analysis (Fig. 4, lane 2). The column flow-through fraction was then analyzed using the 5′ probe to detect spliced, nonpolyadenylated transcripts (Fig. 4, lane 3). This analysis shows that no spliced transcripts are detectable in the poly(A)$^-$ fraction. Although it is possible that these transcripts might be very unstable and difficult to detect, this assay demonstrates that in the steady-state pool, there are no spliced transcripts which have escaped polyadenylation. Nonetheless, we do find polyadenylated transcripts which remain unspliced from this construct while nearly all of the polyadenylated transcripts with the wild-type intron are spliced. Together, these experiments show that substitution with the weak poly(A) site results in lower levels of spliced steady-state RNA. Therefore, we conclude that the strength of the poly(A) site does have an effect on splicing efficiency.

DISCUSSION

The data presented in this report demonstrate the stimulatory effect of an intron on polyadenylation and conversely, the positive role of a poly(A) site on splicing of the terminal intron. In our previous work, we noted a marked stimulation of the HIV-1 poly(A) site, even in competition with a strong L3 poly(A) site downstream, when an intron was placed between the promoter and the poly(A) site (DeZazzo *et al.,* 1992). This enhancement was dependent on the presence of U3 sequences, as stimulation of the relative use of the HIV-1 site was not seen in constructs that lacked U3 sequences. Additionally, we observed an increase in steady-state levels of processed RNA that was independent of the strength of the poly(A) site. In the current study, our objective was to determine if the stimulation caused by the intron was a direct result of the splicing reaction or due to some other enhancement of polyadenylation by the sequences contained within the intron. To inhibit splicing through minimal changes of the intron sequence, we chose to mutate the splice acceptor, based on *in vitro* results that demonstrated that such a mutation completely abolishes splicing (Niwa *et al.,* 1990). We found that this mutation did not abolish splicing *in vivo,* but instead led to activation of a cryptic splice acceptor. Splicing at the cryptic acceptor was as efficient as splicing at the wild-type acceptor when the strong, U3-containing HIV-1 poly(A) site was used. In addition, it had the same stimulatory effect on polyadenylation as the wild-type intron. Mutating the cryptic acceptor led to activation of a second cryptic acceptor farther downstream, again with equivalent effects on polyadenylation (data not shown); therefore we decided to mutate the polypyrimidine tract and branch point residue to abolish splicing. We found that splicing of this mutant intron is completely inhibited *in vivo* and that the observed stimulation of both polyadenylation and steady-state RNA levels is also abrogated. We feel it is unlikely that we have removed any enhancing or accessory sequences directly involved in polyadenylation as a result of these changes and therefore, taken together, the data from the splicing mutants show that it is indeed the splicing of the intron which results in stimulation of the HIV-1 poly(A) site.

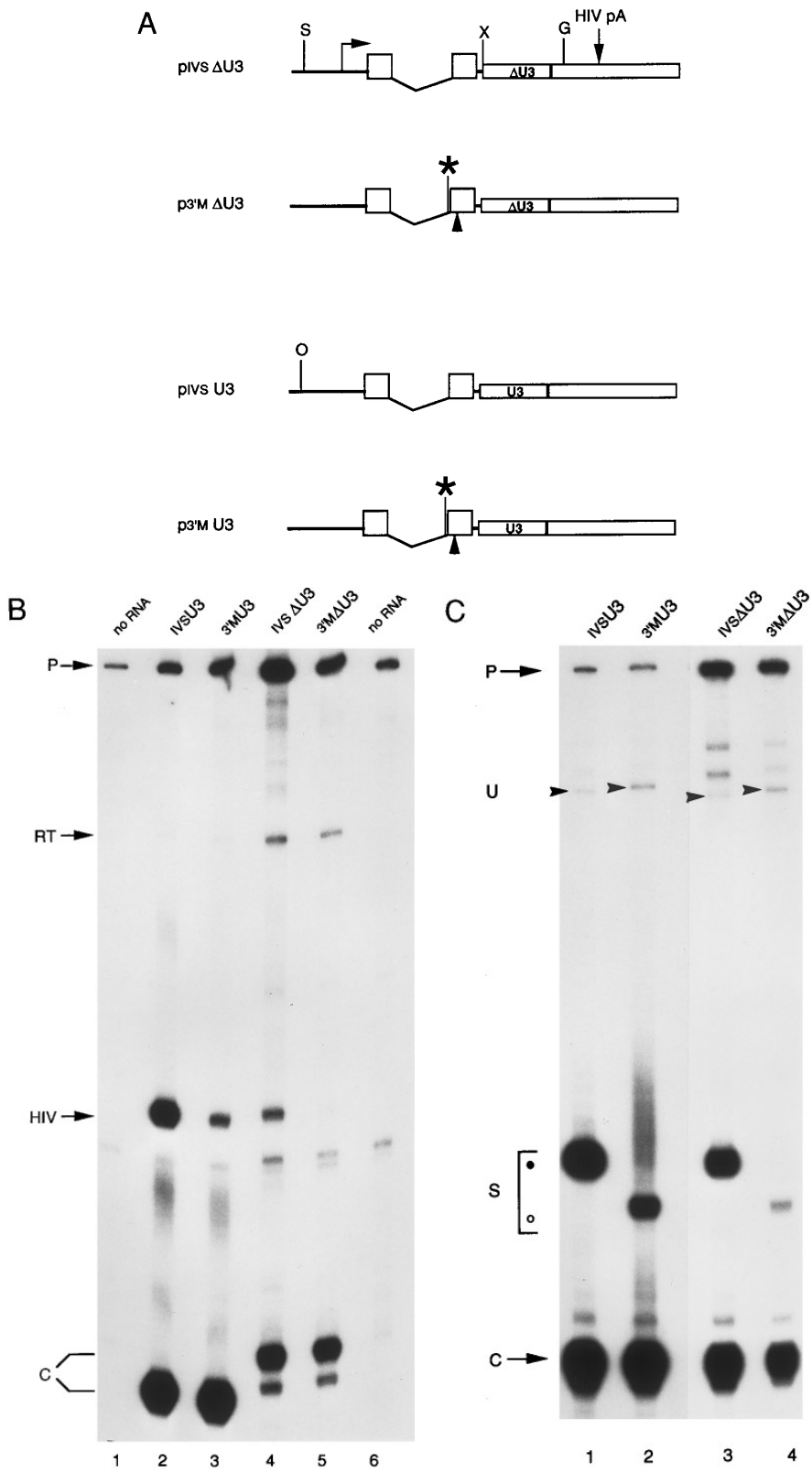

FIG. 3. S1 analysis of polyadenylation and splicing levels in constructs containing weak or strong HIV-1 poly(A) sites. (A) Diagram of plasmids. Single poly(A) site constructs contain either a weak (ΔU3) or a strong (U3) HIV-1 poly(A) site following a wild-type or 3′ mutant intron. Symbols are as in Fig. 1A. S denotes *Stu*I. (B) S1 analysis of 3′ end formation. Lanes 2–5 show S1 products of the poly(A)$^+$ fraction from 50 μg transfection RNA. Lanes 1 and 6 are samples containing probe (U3 or ΔU3, respectively) but no RNA. Full-length probe, HIV-1, and control bands are indicated as in Fig. 2. RT (read through) indicates where the probe sequence diverges from the RNA sequence. (C) S1 analysis of splicing levels in poly(A)$^+$ fraction. Lanes 1 and 2 are from constructs containing the strong HIV-1 poly(A) site downstream of the wild-type or the 3′ mutant intron, respectively. Similarly, lanes 3 and 4 are from the constructs bearing the weak HIV-1 poly(A) site following either of the introns. Unspliced and spliced products are indicated for both introns as in Fig. 2B.

TABLE 2

Splicing Efficiencies

Construct	% Spliced[a]
pIVSU3	99.7 ± 0.1
p3′MU3	94 ± 2
pIVSΔU3	95 ± 2
p3′MΔU3	71 ± 7

[a] Percentage spliced indicates (amount spliced/total poly(A)$^+$ RNA) × 100.

Insertion of a functional intron also increased the steady-state level of poly(A)$^+$ RNA. It is now generally accepted that to achieve high levels of gene expression, one must include an intron in the expression construct. The mechanism of how the intron results in higher expression is not clear, but is likely to involve increased nuclear RNA stability, polyadenylation, and/or nucleocytoplasmic transport of the processed mRNA. In a study addressing this question directly, several steps in mRNA biogenesis were analyzed for two constructs that differed only in the presence or absence of a complete intron (Huang and Gorman, 1990). All parameters tested were equal for the two constructs with the exception of levels of polyadenylated, nuclear RNA, suggesting that splicing may be coupled to the polyadenylation/transport pathway. Other laboratories have also described dramatic differences in expression levels of genes that contain an intron compared to those that lack the intervening sequence (Buchman and Berg, 1988; Ryu and Mertz, 1989; Chiou *et al.,* 1991; Huang and Gorman, 1990). In these reports, splicing was crippled as a result of gross deletions of the intron, and reductions in steady-state RNA levels were observed. The experiments in the present study use defined mutations to abolish splicing without making such drastic alterations to the intron. When splicing was inhibited, the stimulatory effect on the poly(A) site was lost and therefore, our results are consistent with the exon definition model in which splicing and polyadenylation are coupled processing events in intact cells.

An additional objective of this paper was to determine if the poly(A) site was having any effect on the splicing reaction. We reasoned that because splicing is able to play such a dramatic role in the overall biogenesis of mRNA through its effect on 3′ end processing, it is possible that the stimulation might be bidirectional. In our experiments, we used the activation of a cryptic splice acceptor as a sensitive readout for any such stimulation. By deleting the U3 sequences from the HIV-1 poly(A) site, we have converted it to a very weak poly(A) site (DeZazzo *et al.,* 1991; Valsamakis *et al.,* 1991; Brown *et al.,* 1991). Analysis of the splicing efficiencies of these constructs containing wild-type or mutant introns demonstrated that splicing occurs at a lower rate when the poly(A) site is weak. We therefore conclude that poly(A) site strength does have an impact on the splicing of the terminal intron *in vivo.* What effect it may have on splicing of other introns farther upstream has not been determined in these experiments as our constructs contain only one intron. *In vitro* studies have demonstrated that mutation of the AAUAAA sequence to inhibit polyadenylation results in a reduction of splicing of the proximal intron but not the more distal introns, supporting the idea that polyadenylation is coupled only to splicing of the terminal intron (Niwa and Berget, 1991). Although our results are in agreement with these *in vitro* studies, they conflict with a report in which no change in splicing levels was de-

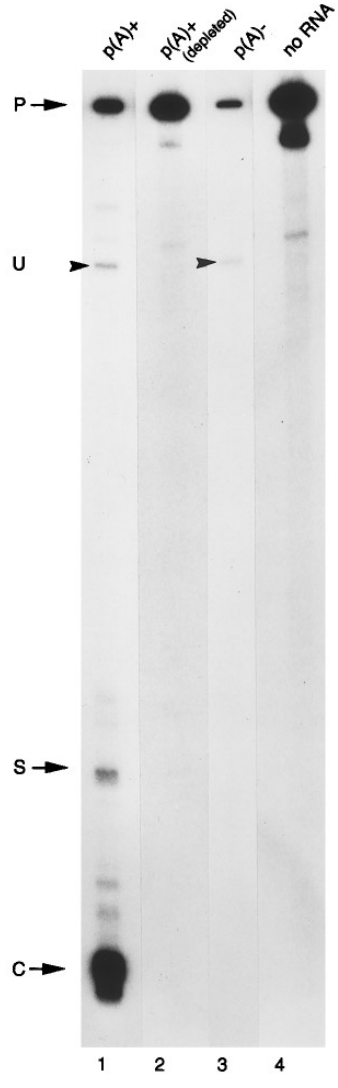

FIG. 4. S1 analysis of poly(A)$^-$ fraction. Total transfection RNA from p3′MΔU3 was oligo(dT) selected until the transfection control signal was no longer detectable in the column eluates with the 5′ end-labeled probe (lane 2). Column flow-through was then analyzed as poly(A)$^-$ with the same probe (lane 3). Unspliced and spliced products are indicated as in Fig. 2B. Lane 1, poly(A)$^+$ fraction, lane 4, no RNA.

tected *in vivo* when the efficiency of 3′ end formation in the triosephosphate isomerase pre-mRNA was altered by deletion or substitution of the poly(A) site (Nesic *et al.,* 1995). These differences may be due to (i) a higher level of sensitivity in our assay which used the cryptic splice acceptor, (ii) a reflection of behavioral differences between viral and cellular processing signals, or (iii) simply an indication that transcription units are regulated in a variety of ways.

An interesting question is what is the mechanism by which these two seemingly independent processing events can stimulate each other? With respect to steady-state RNA levels, it is possible that the assembly of the spliceosome on the transcript could help stabilize the pre-mRNA in the nucleus until processed, followed by the efficient export of the mRNA to the cytoplasm. Perhaps this is how introns can stimulate "generic" gene expression, but in the case of HIV-1, splicing of the intron in the presence of U3 sequences produces an additive effect on polyadenylation (DeZazzo *et al.,* 1992). It has been well documented that U3 sequences are important for efficient 3′ end formation in HIV-1 (DeZazzo *et al.,* 1991; Valsamakis *et al.,* 1991; Brown *et al.,* 1991). Recent studies have shown that these sequences enhance polyadenylation by stabilizing the binding of CPSF at the poly(A) site (Gilmartin *et al.,* 1995). At the 5′ end of the terminal exon, splicing factors, including U1 snRNP, assemble the spliceosome on the terminal intron. A component of the U1 snRNP, the U1A protein, has been shown to also bind to distinct sequences in the upstream efficiency element of SV40 late poly(A) site. This abundant splicing factor can bind simultaneously to the U1 snRNA and the SV40 pre-mRNA (Wassarman and Steitz, 1993; Lutz and Alwine, 1994). Alwine and colleagues have extended their analysis of SV40 pre-mRNA processing to show that the U1A protein directly contacts the 160-kDa protein subunit of the polyadenylation factor CPSF (Lutz *et al.,* 1996). These reports implicate U1A as the molecule which coordinates the polyadenylation and splicing machineries to allow crosstalk to occur over the terminal exon. Taken together, the data suggest a mechanism of cooperation between splicing and polyadenylation factors through U1A and CPSF to define the terminal exon of HIV-1.

At this point, it is important to note that although functional introns can stimulate polyadenylation and the subsequent gene expression, several reports have demonstrated that splicing sequences and factors can also have negative influences on polyadenylation. The U1A protein has been shown to downregulate its own expression by inhibiting polyadenylation of U1A pre-mRNA through a direct interaction with poly(A) polymerase (Gunderson *et al.,* 1994). It has also been shown that the U1 snRNP can interfere with exon definition by binding to a 5′ splice site inserted within a 3′ terminal exon (Niwa *et al.,* 1992). In these experiments, polyadenylation was inhibited through depressed binding of the polyadenylation factor CstF. Perhaps by a similar mechanism, bovine papillomavirus late gene expression is inhibited by a consensus 5′ splice site found in the terminal exon (Furth *et al.,* 1994). This splice site is not used in these transcripts, but its inhibitory effects are dependent on binding to the U1 snRNP. Based on this information, it is possible that the reduction in steady state RNA levels that we observe for the branch point–polypyrimidine tract mutant is due to the fact that the splice donor is not paired with a functional splice acceptor. Consequently, interactions of splicing factors at the 5′ splice site may interfere with terminal exon definition and polyadenylation directly, or simply result in nuclear retention of the transcript. These examples demonstrate the variety of mechanisms by which sequences and processing factors contribute to the regulation of gene expression.

We would argue that in the HIV-1 provirus, these reciprocal effects may be important for the regulation of RNA processing. The 3′ splice sites in HIV-1 are suboptimal sequences, a condition that is important for the balanced production of spliced, partially spliced, and unspliced transcripts required for a productive infection (Dyhr-Mikkelsen and Kjems, 1995; O'Reilly *et al.,* 1995). In addition, terminal sequence redundancy in the provirus results in polyadenylation signals at both the 5′ and the 3′ ends of the pre-mRNA that require regulation: polyadenylation must occur at the site in the 3′ LTR to ensure adequate gene expression and genome production. As mentioned above, studies from our laboratory and others have demonstrated a significant role for U3 sequences in stimulating the 3′ poly(A) site over the 5′ poly(A) site (DeZazzo *et al.,* 1991, 1992; Brown *et al.,* 1991; Valsamakis *et al.,* 1991). In a recent report examining the mechanism of this poly(A) site competition, it was suggested that U3 sequences function to strengthen the HIV-1 poly(A) site in the event that it must compete with downstream processing signals in the cellular genome and that U3 sequences are not necessary for polyadenylation nor do they enhance polyadenylation when there is no nearby competing site (Ashe *et al.,* 1995). We propose that, in addition to the demonstrated importance of U3 sequences for efficient polyadenylation, the positive effect that splicing has on the U3-containing HIV-1 poly(A) site is an extra layer of insurance that proviral transcripts will be polyadenylated properly, leading to appropriate gene expression. Conversely, the reciprocal stimulation of splicing by polyadenylation ensures that HIV-1 transcripts will become fully spliced during the early phase of the infection. This results in efficient expression of mRNAs encoding transactivating and regulatory proteins, such as Tat and Rev, which are needed in sufficient quantity to regulate further mRNA biogenesis (Cullen, 1992). Later in the infection, Rev functions to regulate the

splicing and/or transport of singly spliced messages that encode the envelope protein and of unspliced transcripts which encode the gag–pol precursor protein or serve as progeny viral genomes. Thus, it appears that HIV-1 has taken advantage of the many nuances RNA processing affords to control gene regulation for optimal virion production.

ACKNOWLEDGMENTS

We thank the members of the Imperiale laboratory for critical comments and suggestions, David Karow for technical assistance, and Sue Berget for providing the splicing cassette plasmids pMXSVL and pMXSVL 3′. This work was supported by GM34902 from the National Institutes of Health, and by RP00042 for computing resources. J.M.S. was supported in part by a Horace Rackham Predoctoral Fellowship from the University of Michigan.

REFERENCES

Ashe, M. P., Griffin, P., James, W., and Proudfoot, N. J. (1995). Poly(A) site selection in the HIV-1 provirus: Inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev.* 9, 3008–3025.

Brown, P. H., Tiley, L. S., and Cullen, B. R. (1991). Efficient polyadenylation within the human immunodeficiency virus type 1 long terminal repeat requires flanking U3-specific sequences. *J. Virol.* 65, 3340–3343.

Buchman, A. R., and Berg, P. (1988). Comparison of intron-dependent and intron-independent gene expression. *Mol. Cell. Biol.* 8, 4395–4405.

Carswell, S., and Alwine, J. C. (1989). Efficiency of utilization of the simian virus 40 late polyadenylation site: Effects of upstream sequences. *Mol. Cell. Biol.* 9, 4248–4258.

Chiou, H. C., Dabrowski, C., and Alwine, J. C. (1991). Simian virus 40 late mRNA leader sequences involved in augmenting mRNA accumulation via multiple mechanisms, including increased polyadenylation efficiency. *J. Virol.* 65, 6677–6685.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* 162, 156–159.

Cullen, B. R. (1992). Mechanism of action of regulatory proteins encoded by complex retroviruses. *Microbiol. Rev.* 56(3), 375–394.

DeZazzo, J. D., Kilpatrick, J. E., and Imperiale, M. J. (1991). Involvement of long terminal repeat U3 sequences overlapping the transcription control region in human immunodeficiency virus type 1 mRNA 3′ end formation. *Mol. Cell. Biol.* 11, 1624–1630.

DeZazzo, J. D., Scott, J. M., and Imperiale, M. J. (1992). Relative roles of signals upstream of AAUAAA and promoter proximity in regulation of human immunodeficiency virus type 1 mRNA 3′ end formation. *Mol. Cell. Biol.* 12, 5555–5562.

DeZazzo, J. D., and Imperiale, M. J. (1989). Sequences upstream of AAUAAA influence poly(A) site selection in a complex transcription unit. *Mol. Cell. Biol.* 9, 4951–4961.

Dougherty, J. P., and Temin, H. M. (1987). A promoterless retroviral vector indicates that there are sequences in U3 required for 3′ RNA processing. *Proc. Natl. Acad. Sci. USA* 84, 1197–1201.

Dyhr-Mikkelsen, H., and Kjems, J. (1995). Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript in vitro. *J. Biol. Chem.* 270, 24060–24066.

Furth, P. A., Choe, W., Rex, J. H., Byrne, J. C., and Baker, C. C. (1994). Sequences homologous to 5′ splice sites are required for the inhibitory activity of papillomavirus late 3′ untranslated regions. *Mol. Cell. Biol.* 14(8), 5278–5289.

Gilmartin, G. M., Fleming, E. S., Oetjen, J., and Graveley, B. R. (1995). CPSF recognition of an HIV-1 mRNA 3′-processing enhancer: Multiple sequence contacts involved in poly(A) site definition. *Genes Dev.* 9, 72–83.

Graham, F. L., Smiley, J., Russell, W. C., and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* 36, 59–74.

Green, M. R. (1991). Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu. Rev. Cell Biol.* 7, 559–599.

Gunderson, S. I., Beyer, K., Martin, G., Boelens, W. C., and Mattaj, I. W. (1994). The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell* 76, 531–541.

Huang, M. T., and Gorman, C. M. (1990). Intervening sequences increase efficiency of RNA 3′ processing and accumulation of cytoplasmic RNA. *Nucleic Acids Res.* 18, 937–947.

Humphrey, M. B., Bryan, J., Cooper, T. A., and Berget, S. M. (1995). A 32-nucleotide exon-splicing enhancer regulates usage of competing 5′ splice sites in a differential internal exon. *Mol. Cell. Biol.* 15(8), 3979–3988.

Lavigueur, A., La Branche, H., Kornblihtt, A. R., and Chabot, B. (1993). A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes Dev.* 7, 2405–2417.

Liu, X., and Mertz, J. E. (1993). Polyadenylation site selection cannot occur in vivo after excision of the 3′-terminal intron. *Nucleic Acids Res.* 21, 5256–5263.

Lou, H., Gagel, R. F., and Berget, S. M. (1996). An intron enhancer recognized by splicing factors activates polyadenylation. *Genes Dev.* 10, 208–219.

Lutz, C. S., and Alwine, J. C. (1994). Direct interaction of the U1snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes Dev.* 8, 576–586.

Lutz, C. S., Murthy, K. G. K., Schek, N., O'Connor, J. P., Manley, J. L., and Alwine, J. C. (1996). Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes Dev.* 10, 325–337.

Manley, J. L. (1988). Polyadenylation of mRNA precursors. *Biochim. Biophys. Acta* 950, 1–12.

Miller, J. T., and Stoltzfus, C. M. (1992). Two distant upstream regions containing cis-acting signals regulating splicing facilitate 3′-end processing of avian sarcoma virus RNA. *J. Virol.* 66, 4242–4251.

Moore, M. J., Query, C. C., and Sharp, P. A. (1993). *In* "The RNA World" (R. F. Gesteland and J. F. Atkins, Eds.), pp. 303–357. Cold Spring Harbor Laboratory Press, Plainview, NY.

Nesic, D., Cheng, J., and Maquat, L. E. (1993). Sequences within the last intron function in RNA 3′-end formation in cultured cells. *Mol. Cell. Biol.* 13, 3359–3369.

Nesic, D., and Maquat, L. E. (1994). Upstream introns influence the efficiency of final intron removal and RNA 3′-end formation. *Genes Dev.* 8, 363–375.

Nesic, D., Zhang, J., and Maquat, L. E. (1995). Lack of an effect of the efficiency of RNA 3′-end formation on the efficiency of removal of either the final or the penultimate intron in intact cells. *Mol. Cell. Biol.* 15, 488–496.

Niwa, M., Rose, S. D., and Berget, S. M. (1990). In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev.* 4, 1552–1559.

Niwa, M., and Berget, S. M. (1991). Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev.* 5, 2086–2095.

Niwa, M., MacDonald, C. C., and Berget, S. M. (1992). Are vertebrate exons scanned during splice-site selection? *Nature* 360, 277–280.

O'Reilly, M. M., McNally, M. T., and Beemon, K. L.(1995). Two strong 5′ splice sites and competing, suboptimal 3′ splice sites involved in

alternative splicing of human immunodeficiency virus type 1 RNA. *Virology* 213, 373–385.

Reed, R. (1989). The organization of 3′ splice-site sequences in mammalian introns. *Genes Dev.* 3, 2113–2123.

Robberson, B. L., Cote, G. J., and Berget, S. M. (1990). Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* 10, 84–94.

Roscigno, R. F., Weiner, M., and Garcia-Blanco, M. A. (1993). A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing. *J. Biol. Chem.* 268, 11222–11229.

Russnak, R., and Ganem, D. (1990). Sequences 5′ to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. *Genes Dev.* 4, 764–776.

Ryu, W. S., and Mertz, J. E. (1989). Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm. *J. Virol.* 63, 4386–4394.

Sanfacon, H., Brodmann, P., and Hohn, T. (1991). A dissection of the cauliflower mosaic virus polyadenylation signal. *Genes Dev.* 5, 141–149.

Tian, M., and Maniatis, T. (1993). A splicing enhancer complex controls alternative splicing of doublesex pre-mRNA. *Cell* 74, 105–114.

Tian, M., and Maniatis, T. (1994). A splicing enhancer exhibits both constitutive and regulated activities. *Genes Dev.* 8, 1703–1712.

Valsamakis, A., Zeichner, S., Carswell, S., and Alwine, J. C. (1991). The human immunodeficiency virus type 1 polyadenylylation signal: A 3′ long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylylation. *Proc. Natl. Acad. Sci. USA* 88, 2108–2112.

Villarreal, L. P., and White, R. T. (1983). A splice junction deletion deficient in the transport of RNA does not polyadenylate nuclear RNA. *Mol. Cell. Biol.* 3, 1381–1388.

Wahle, E. (1995). 3′-end cleavage and polyadenylation of mRNA precursors. *Biochim. Biophys. Acta* 1261, 183–194.

Wahle, E., and Keller, W. (1992). The biochemistry of 3′-end cleavage and polyadenylation of messenger RNA precursors. *Annu. Rev. Biochem.* 61, 419–440.

Wassarman, K. M., and Steitz, J. A. (1993). Association with terminal exons in pre-mRNAs: A new role for the U1 snRNP? *Genes Dev.* 7, 647–659.

Watakabe, A., Tanaka, K., and Shimura, Y. (1993). The role of exon sequences in splice site selection. *Genes Dev.* 7, 407–418.